

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
27 January 2005 (27.01.2005)

PCT

(10) International Publication Number
WO 2005/008525 A1

(51) International Patent Classification⁷: **G06F 17/30**

(74) Agent: **GROENENDAAL, Antonius, W., M.**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(21) International Application Number:
PCT/IB2004/051244

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 16 July 2004 (16.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
03102240.3 21 July 2003 (21.07.2003) EP

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

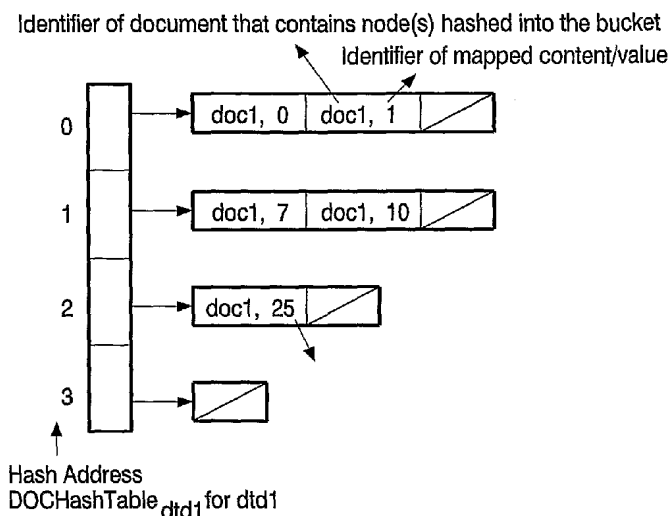
(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **JONKER, Willem** [NL/FR]; c/o Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **FENG, Ling** [CN/NL]; c/o Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

Published:
— with international search report

[Continued on next page]

(54) Title: METHOD OF SEARCHING IN A COLLECTION OF DOCUMENTS



(57) **Abstract:** The present invention relates to a method of searching in a collection of documents having a tree-like structure and complying with at least one document structure definition in a collection of document structure definitions. This method comprises the steps of receiving (1) a certain branch, determining (3) a subset of the collection of document structure definitions which allow the certain branch to exist in a document complying to the document structure definition, determining (5) a subset of the collection of documents which comprises all documents of the collection of documents complying to any one of the document structure definitions in the subset, and searching (7) for at least part of the certain branch in each document. The present invention further relates to a method of indexing a collection of documents for facilitating the method of searching in a collection of documents of the invention. The present invention also relates to computer program products enabling a programmable device to carry out the methods of the invention and to electronic devices for carrying out the methods of the invention.

WO 2005/008525 A1



— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.